



Image moments-based ultrasound visual servoing

R. Mebarki, A. Krupa, François Chaumette

► To cite this version:

R. Mebarki, A. Krupa, François Chaumette. Image moments-based ultrasound visual servoing. IEEE Int. Conf. on Robotics and Automation, ICRA'08, 2008, Pasadena, California, France. pp.113-119. inria-00351872

HAL Id: inria-00351872

<https://inria.hal.science/inria-00351872>

Submitted on 12 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Image Moments-based Ultrasound Visual Servoing

Rafik Mebarki, Alexandre Krupa and François Chaumette

Abstract—A new visual servoing method based on B-mode ultrasound images is proposed to automatically control the motion of a 2D ultrasound probe held by a medical robot in order to reach a desired B-scan image of an object of interest. In this approach, combinations of image moments extracted from the current observed object cross-section are used as feedback visual features. The analytical form of the interaction matrix, relating the time variation of these visual features to the probe velocity, is derived and used in the control law. Simulations performed with a static ultrasound volume containing an egg-shaped object, and *in-vitro* experiments using a robotized ultrasound probe that interacts with a rabbit heart immersed in water, show the validity of this new approach and its robustness with respect to modeling and measurements errors.

Index Terms—Visual servoing, ultrasound, image moments, medical robotics.

I. INTRODUCTION

In order to assist radiologist diagnostics or ultrasound guided interventions, we propose to automatically position an ultrasound (US) probe held by a medical robot in such a way to reach and track an appropriate US image cross-section of a given observed object. This will allow different kinds of application. For example for pathology analysis, it can be helpful to automatically and accurately position the US probe in order to obtain a 2D image cross-section of a tumor having a maximum similarity with one derived from a pre-operative 3D imaging that was performed with the same (US) or other imaging modality (MRI, CT-SCAN). Autonomously reaching and maintaining an appropriate tumor cross-section would also help the surgeon to perform needle insertion during a biopsy or a radio-frequency ablation procedure. Towards that goal, we propose a new visual servoing method based on ultrasound images and specially on the use of image moments to control a medical robot holding an US probe.

Up until now, only few research works have focused on visual servoing based on ultrasound images. In [1], a robot-assisted system designed for 3D ultrasound imaging [2] is controlled by ultrasound visual servoing in order to center the section of an artery within the 2D US image. However, only the in-plane motions (3 DOF) of the probe corresponding to the two translations and the rotation in the US image plane are controlled directly from the image, while the other 3 DOF are teleoperated by the user. In [3], the development of a real-time ultrasound-guided needle insertion medical robot for per-cutaneous cholecystostomy has been presented. However, only two of the 5 DOF of the robot, which was specifically

designed for per-cutaneous ultrasound-guided therapy, are controlled to position the needle which is mechanically constrained to lie in the motionless probe observation plane. In these two prior works, visual servoing was performed thanks to a simple model of the probe interaction that does not require any 3D information for in-plane motions. However, if the visual task consists also in controlling the out-of-plane motions these previous methods can not be used and it is necessary to model the complete interaction between the probe and the object. In fact an ultrasonic probe provides full information only in its observation plane whereas a camera provides a projection of the 3D world to a 2D image. Another alternative is the use of a 3D US probe as in [4] where a surgical instrument is controlled using position-based servoing. However, for the moment 3D US probes are expensive and provide only small volume with low voxel resolution. Therefore, in what follows we will only consider the use of standard 2D probes, which are more widespread in medical centers.

Some recent studies eliminate the requirement of controlling by visual servoing only the DOF contained in the ultrasound plane by modeling the interaction between the ultrasound probe and a given observed object. In [6] and [7] two image points corresponding to the intersections of a laparoscopic instrument forceps with the ultrasound plane are used to servo the 4 DOF of the instrument.

In [8] visual features corresponding to the intersection points of a cross-wire object with the US plane allow to automatically perform the calibration of a robotized 3D ultrasound imaging system. In [9], a first ultrasound image-based visual servoing that allows to control the motion of an US robotized probe in order to reach a desired image section of an egg-shaped object has been presented. In this previous work, the image edge was modelled by a third degree polynomial whose coefficients were selected as feedback features in the control law. However polynomial coefficients have no physical significations, they are very sensitive to image noise and limit the use of the method to specific object shapes.

In this paper, we develop a new visual servoing method based on image moments. These features are generic with intuitive and geometric meaning and are robust with respect to measurement perturbations, allowing, thus, a robust control system. Indeed, they have been widely used in computer vision for a very long time, especially for pattern recognition applications [10] and have been recently introduced in visual servoing using the perspective projection [11]. It seems thus interesting to use them as feedback information in US visual servoing where very noisy images are considered. However

The authors are with IRISA, INRIA Rennes-Bretagne Atlantique, Campus de Beaulieu, 35 042 Rennes-cedex, France {Rafik.Mebarki, Alexandre.Krupa, Francois.Chaumette}@irisa.fr

for ultrasound imaging, the modeling part differs from the case of camera perception and has to be reformulated.

The paper is organized as follows: In the next section, the analytical form of the interaction matrix, relating the time variation of the US image moments to the velocity, is derived. Then, in section III, a set of combinations of moments is selected as feedback information and the control law is derived. Simulations and experimental results are presented and discussed in Section IV. They demonstrate the validity of the proposed method. Finally, concluding remarks are given in Section V.

II. MODELING

The robotic task consists in automatically positioning an US probe held by a medical robot arm in such a way to view a desired cross-section of a given object as depicted in Fig. 1. To design the visual servoing control scheme it is essential to choose appropriate visual features and determine the interaction matrix that relates their variation to the probe velocity. This modeling aspect is the aim of this section.

A. Image Moments Interaction Matrix Modeling

Let \mathcal{O} be the object referring to the organ with which the US probe is interacting. The moments m_{ij} of order $i + j$ are defined by:

$$m_{ij} = \iint_{\mathcal{S}} f(x, y) dx dy \quad (1)$$

where $f(x, y) = x^i y^j$ and (x, y) represent US image point coordinates. \mathcal{S} is the US image cross-section resulting from the intersection of the object \mathcal{O} with the US probe plane (see Fig. 1). It would be also possible to generalize f to $f(x, y) = x^i y^j g(x, y)$ where $g(x, y)$ is associated to the gray level of each image pixel. However that would perturb considerably the control system since the US images are very noisy and present artifacts. The objective is to determine the analytical form of the time variation \dot{m}_{ij} of moments m_{ij} in function of the probe velocity $\mathbf{v} = (v \ \omega)$ such as:

$$\dot{m}_{ij} = \mathbf{L}_{m_{ij}} \mathbf{v} \quad (2)$$

where (v_x, v_y, v_z) and $\omega = (\omega_x, \omega_y, \omega_z)$ represent respectively the translational and the rotational velocity components and $\mathbf{L}_{m_{ij}}$ is the interaction matrix related to m_{ij} denoted by:

$$\mathbf{L}_{m_{ij}} = \begin{bmatrix} m_{vx} & m_{vy} & m_{vz} & m_{wx} & m_{wy} & m_{wz} \end{bmatrix} \quad (3)$$

The time variation of moments is given by [11]:

$$\dot{m}_{ij} = \iint_{\mathcal{S}} \left[\frac{\partial f}{\partial x} \dot{x} + \frac{\partial f}{\partial y} \dot{y} + f(x, y) \left(\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} \right) \right] dx dy \quad (4)$$

where (\dot{x}, \dot{y}) is the velocity of an image point (x, y) belonging to the section \mathcal{S} . So, in order to determine the relation giving \dot{m}_{ij} in function of \mathbf{v} , the image point velocity (\dot{x}, \dot{y}) has to be expressed in function of \mathbf{v} . To determine the components of $\mathbf{L}_{m_{ij}}$, the two kinds of probe motion which are the in-plane and out-of-plane motions are considered below.

1) *In-plane motions*: The probe, in this case, remains in its initial plane. For in-plane motions, the velocity ${}^s\dot{\mathbf{P}}$ of any point \mathbf{P} of \mathcal{O} in \mathcal{S} with respect to the probe Cartesian frame $\{R_s\}$ can be expressed using the fundamental kinematic relationship:

$${}^s\dot{\mathbf{P}} = \begin{bmatrix} -\mathbf{I}_3 & [{}^s\mathbf{P}]_{\times} \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \quad (5)$$

where ${}^s\mathbf{P} = (x, y, 0)$ is the position of the point \mathbf{P} expressed in the US probe frame $\{R_s\}$. Since the probe remains in the same plane, only the three in-plane components v_x, v_y and ω_z of \mathbf{v} are considered. Therefore ${}^s\dot{\mathbf{P}}$ which is nothing but equal to $(\dot{x}, \dot{y}, 0)$ is given by:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -1 & 0 & y \\ 0 & -1 & -x \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ w_z \end{bmatrix} \quad (6)$$

from which we deduce: $\frac{\partial \dot{x}}{\partial x} = 0, \frac{\partial \dot{y}}{\partial y} = 0$. Substituting (6) into (4) and remembering that $f(x, y) = x^i y^j$, $\partial f / \partial x = i x^{i-1} y^j$ and $\partial f / \partial y = j x^i y^{j-1}$, the three coefficients m_{vx} , m_{vy} and m_{wz} of the interaction matrix (3) relating to in-plane probe velocity components are derived as follows:

$$\begin{cases} m_{vx} &= -i m_{i-1,j} \\ m_{vy} &= -j m_{i,j-1} \\ m_{wz} &= i m_{i-1,j+1} - j m_{i+1,j-1} \end{cases} \quad (7)$$

The time variation of moment of order $i + j$, when the US probe lies in its initial plane, can thus be easily computed from moments of order $i + j - 1$ and $i + j$.

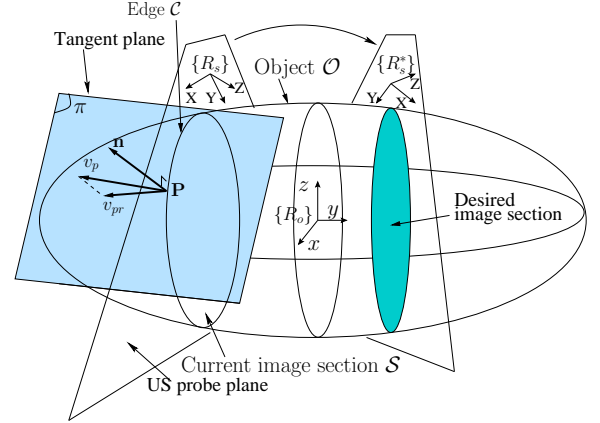


Fig. 1. Interaction between the ultrasound probe and the object

Now, we consider several classical features: the section area a , the center of mass coordinates of the image section x_g, y_g and its main orientation α . These are defined in terms of the image moments as follows:

$$\begin{cases} a &= m_{00} \\ x_g &= m_{10}/a \\ y_g &= m_{01}/a \\ \alpha &= \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} + \mu_{02}} \right) \end{cases} \quad (8)$$

where $\mu_{11} = m_{11} - a x_g y_g$, $\mu_{20} = m_{20} - a x_g^2$ and $\mu_{02} = m_{02} - a y_g^2$ are the centered moments of order 2. Using (7),

we can easily relate their time variation to the in-plane probe velocity and obtain:

$$\begin{bmatrix} \dot{a} \\ \dot{x}_g \\ \dot{y}_g \\ \dot{\alpha} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & y_g \\ 0 & -1 & -x_g \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ w_z \end{bmatrix} \quad (9)$$

We can notice that the section area a is invariant to in-plane motions of the probe and therefore seems to be a good feature for out-of-plane motion control whereas x_g , y_g and α are well adapted for in-plane motion control, with a good decoupling property that can be seen by referring to the triangular part of the matrix in (9).

2) *Out-of-plane motions:* When the US probe gets out of its initial plane, the image variation of the edge \mathcal{C} of \mathcal{S} depends strongly on the 3D shape of the object. The reason is that image points do not correspond to the same 3D object points when out-of-plane motions are applied. In this case the edge points resulting from the intersection between the ultrasound plane and the object surface are affected in the same manner as if they slide on the object surface (see Fig 2). Therefore we can consider that an edge point \mathbf{P} is a moving point $\mathbf{P}(t)$ that is constrained to remain in the image plane while sliding on the object surface. That means that its velocity with respect to the object frame $\{R_o\}$ expressed in the probe frame $\{R_s\}$ depends both on the probe out-of plane velocity and the surface shape. However, its z-axis velocity component is related only to the probe velocity because the point is constrained to remain in the image plane and is given simply by:

$$v_{pz} = v_z + y \omega_x - x \omega_y \quad (10)$$

where v_z , ω_x and ω_y are the three out-of-plane velocity components of the probe. The objective consists then in determining the x and y axis velocity components that represent directly the image variation of the point. To determine them, we consider that the object surface can be locally approximated by a tangent plane π as shown in Fig. 1, and project on it the velocity component v_{pz} as depicted in 2D in Fig. 2. This projection is performed by the following

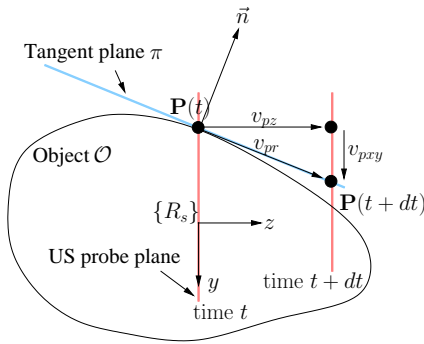


Fig. 2. Tangent plane used to determine the image point velocity

approximation:

$$\mathbf{v}_{pr} \approx \mathbf{n} \times (\mathbf{v}_p \times \mathbf{n}) \quad (11)$$

with $\mathbf{v}_p = [0 \ 0 \ v_{pz}]^T$ and where $\mathbf{n} = [n_x \ n_y \ n_z]^T$ is the unitary normal vector to π expressed in $\{R_s\}$. Developing (11) gives:

$$\mathbf{v}_{pr} = v_{pz} \begin{bmatrix} -n_x n_z \\ -n_y n_z \\ n_x^2 + n_y^2 \end{bmatrix} \quad (12)$$

Substituting v_{pz} from (10) in (12) and projecting \mathbf{v}_{pr} on the US plane (x,y) gives $\mathbf{v}_{pxy} = [\dot{x} \ \dot{y} \ 0]^T$, and consequently the image point variation in function of the out-of-plane probe velocity components as follows:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -n_x n_z & -n_x n_z y & n_x n_z x \\ -n_y n_z & -n_y n_z y & n_y n_z x \end{bmatrix} \begin{bmatrix} v_z \\ w_x \\ w_y \end{bmatrix} \quad (13)$$

As we can see, knowledge of the orientation of the tangent plane π at surface point \mathbf{P} is crucial to relate the image point variation to the out-of-plane probe velocity. Therefore we will now determine the expression of the normal vector \mathbf{n} for a given object. If the object surface can be implicitly defined as a set of points $(^o x, ^o y, ^o z)$ satisfying $F(^o x, ^o y, ^o z) = 0$, then, the normal vector to that surface at a point $(^o x, ^o y, ^o z)$ is given directly by the gradient:

$$^o \mathbf{n} = \frac{\nabla F}{\|\nabla F\|} \quad (14)$$

with $\nabla F = [\partial F / \partial ^o x \ \partial F / \partial ^o y \ \partial F / \partial ^o z]^T$ and where $^o \mathbf{P} = (^o x, ^o y, ^o z)$ is the vector coordinates of a surface point \mathbf{P} expressed in the object Cartesian frame $\{R_o\}$. In the following, we consider the case where the object has an ellipsoidal shape. The choice of this model is motivated by the likeness of this form to usual tumors. Nevertheless our method is generic and can be extended to more complex shapes if the normal vector of the surface can be analytically formulated.

The ellipsoid surface is given by the equation:

$$F(^o x, ^o y, ^o z) = \left(\frac{^o x}{a_1}\right)^2 + \left(\frac{^o y}{a_2}\right)^2 + \left(\frac{^o z}{a_3}\right)^2 - 1 = 0 \quad (15)$$

where a_1, a_2, a_3 are the half length values of the object main axes. Applying (14) gives thus the following normal vector expressed in the object frame $\{R_o\}$ for a given surface point:

$$^o \mathbf{n} = \frac{[\ ^o x/a_1^2 \ ^o y/a_2^2 \ ^o z/a_3^2]^T}{\| [\ ^o x/a_1^2 \ ^o y/a_2^2 \ ^o z/a_3^2] \|} \quad (16)$$

In what follows we approximate $^o \mathbf{n}$ to the expression:

$$^o \tilde{\mathbf{n}} = r \begin{bmatrix} ^o x/a_1^2 \\ ^o y/a_2^2 \\ ^o z/a_3^2 \end{bmatrix} \quad (17)$$

where $r = (a_1 + a_2 + a_3)/3$ is a constant scalar. This approximation does not affect the direction of the normal vector but only its norm. It is performed in order to obtain adequate linear relations, as it will be shown later. This allows to rewrite $^o \tilde{\mathbf{n}}$ in the following linear form:

$$^o \tilde{\mathbf{n}} = r \begin{bmatrix} 1/a_1^2 & 0 & 0 \\ 0 & 1/a_2^2 & 0 \\ 0 & 0 & 1/a_3^2 \end{bmatrix} \begin{bmatrix} ^o x \\ ^o y \\ ^o z \end{bmatrix} = \mathbf{C} \ ^o \mathbf{P} \quad (18)$$

and its expression in the US probe frame $\{R_s\}$ becomes:

$$\tilde{\mathbf{n}} = {}^s\tilde{\mathbf{n}} = {}^s\mathbf{R}_o \circ \tilde{\mathbf{n}} = {}^s\mathbf{R}_o \mathbf{C} \circ \mathbf{P} \quad (19)$$

where ${}^s\mathbf{R}_o$ is the rotation matrix defining the orientation of the object frame $\{R_o\}$ with respect to the probe frame $\{R_s\}$. The vector $\tilde{\mathbf{n}}$ can then be related to the point coordinates ${}^s\mathbf{P} = (x, y, 0)$ expressed in the probe frame by the following expression:

$$\tilde{\mathbf{n}} = {}^s\mathbf{R}_o \mathbf{C} {}^s\mathbf{R}_o^\top ({}^s\mathbf{P} - {}^s\mathbf{t}_o) \quad (20)$$

where ${}^s\mathbf{t}_o$ is the translation defining the position of the object frame origin with respect to the probe frame. Here $\tilde{\mathbf{n}}$ is function of: the image coordinates (x, y) of the considered edge \mathcal{C} point, the pose of the object with respect to the probe frame and the 3D parameters corresponding to the length of the object axes. Developing (20) we obtain the components of the vector $\tilde{\mathbf{n}}$ and therefore the following approximation of the terms required in (13):

$$\begin{cases} \tilde{n}_x \tilde{n}_z = -A_1 x^2 - A_2 x y - A_3 y^2 - A_4 x \\ \quad - A_5 y - A_6 \\ \tilde{n}_y \tilde{n}_z = -B_1 x^2 - B_2 x y - B_3 y^2 - B_4 x \\ \quad - B_5 y - B_6 \end{cases} \quad (21)$$

where $A_{i|i=1..6} = A_i({}^s\mathbf{M}_o, a_{k|k=1..3})$, $B_{i|i=1..6} = B_i({}^s\mathbf{M}_o, a_{k|k=1..3})$ are terms depending on the size of the ellipsoid and the relative pose between frames $\{R_s\}$ and $\{R_o\}$. This relative position which is described by the homogeneous matrix ${}^s\mathbf{M}_o$ will be estimated in Section III-B. Detailed expressions are not given for a lack of place¹. Note that in the simple case of a spherical object, the relation (20) gives $\tilde{\mathbf{n}} = \mathbf{n} = r ({}^s\mathbf{P} - {}^s\mathbf{t}_o)$ and therefore $\tilde{\mathbf{n}}$ does not depend on the rotation matrix between the object and the US probe frame. This is justified by the fact that a sphere does not have any orientation in the 3D space.

By substituting \dot{x} and \dot{y} by (13) in (4) and using (21), we obtain after some simple calculus the expression of \dot{m}_{ij} in function of the out-of-plane probe velocity components v_z , ω_x and ω_y . The remaining components of the interaction matrix $\mathbf{L}_{m_{ij}}$ are thus given by:

$$\begin{cases} m_{vz} = f_{m_{ij}} \\ m_{wx} = f_{m_{i,j+1}} \\ m_{wy} = -f_{m_{i+1,j}} \end{cases} \quad (22)$$

with:

$$\begin{aligned} f_{m_{ij}} = & i (A_6 m_{i-1,j} + A_5 m_{i-1,j+1} + A_3 m_{i-1,j+2}) \\ & + j (B_6 m_{i,j-1} + B_4 m_{i+1,j-1} + B_1 m_{i+2,j-1}) \\ & + (B_2 (j+1) + A_1 (i+2)) m_{i+1,j} \\ & + (B_3 (j+2) + A_2 (i+1)) m_{i,j+1} \\ & + (B_5 (j+1) + A_4 (i+1)) m_{ij} \end{aligned} \quad (23)$$

The time variation of moments of order $i+j$ produced by out-of-plane probe motions can thus be expressed directly from moments of order $i+j-1$ up to $i+j+1$ and the 3D parameters $A_{k|k=1..6}$ and $B_{k|k=1..6}$ related to the

object. It should be noted that $f_{m_{ij}}$ is given here as a linear combination of image moments. This would not be the case if the approximation (17) was not performed. In fact, replacing the constant scalar r by the exact expression that allows to normalize $\tilde{\mathbf{n}}$ such that $\tilde{\mathbf{n}} = \mathbf{n}$ would lead to a more complex relation than (23) including non-linear combinations of moments. In spite of this approximation and the approximation in (11), simulations and experiments presented in Section IV demonstrate the validity of the modeling thanks to the robustness of the visual servoing.

III. VISUAL SERVOING

A. Selection of features for visual servoing

As an ellipsoidal object has been considered in the modeling, the image section provided by the US probe corresponds in all situations to a 2D ellipse. Therefore only 5 visual features are required to describe the object section. This also means that there are an infinity of 3D poses of the probe that give the same image section of the ellipsoidal object. It is of course possible to overcome this limitation by considering more complex and non symmetrical object models, as we plan to do in future works.

For the ellipsoidal object, 5 visual features have to be selected as feedback information in the control scheme. As we mentioned in Section II-A.1, the center of mass coordinates (x_g, y_g) of the image section and its main orientation α are 3 features well adapted for in-plane motion control due to their high dependence on this kind of motion and the significant decoupling between them. Since the surface area a is invariant to in-plane motions, it is consequently an appropriate visual feature for out-of-plane motion control. Concerning the fifth visual feature we propose to use the length of the section's major axis. The visual features vector is thus given by:

$$\mathbf{s} = (x_g, y_g, \alpha, l_1, a) \quad (24)$$

The expression of x_g, y_g, α, a in function of images moments was given in Section II-A.1 and l_1 is defined by:

$$l_1 = \frac{2}{a} \left(\mu_{02} + \mu_{20} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \right) \quad (25)$$

where μ_{11}, μ_{02} and μ_{20} are the centered moments of order 2 defined in Section II-A.1.

The time variation of the visual features vector in function of the probe velocity is written as follows:

$$\dot{\mathbf{s}} = \mathbf{L}_s \mathbf{v} \quad (26)$$

where \mathbf{L}_s is the interaction matrix related to \mathbf{s} that we easily determined from (7) and (22) respectively for the in-plane and out-of-plane velocity components. Note that the fifth feature l_1 was chosen since, in comparison with other possible features that we considered, as for example the ratio on the two main axes, it provides a better conditioning of the interaction matrix \mathbf{L}_s .

¹They can be found on <http://www.irisa.fr/lagadic/team/Rafik.Mebarki-eng.html>

B. Object pose and 3D parameters estimation

The pose between the object and the US frame which is required for on-line updating the interaction matrix is estimated as follows:

$${}^s\mathbf{M}_o = ({}^b\mathbf{M}_s)^{-1} {}^b\mathbf{M}_o \quad (27)$$

where ${}^b\mathbf{M}_s = {}^b\mathbf{M}_e {}^e\mathbf{M}_s$ is the pose of the probe frame expressed in the robot base frame $\{R_b\}$, ${}^b\mathbf{M}_e$ is the robot end-effector pose with respect to the robot base frame given by the forward robot kinematics and ${}^e\mathbf{M}_s$ is a constant homogeneous matrix defining the relative position between the end-effector and the US probe frame. This last matrix contains the spatial parameters of the ultrasound imaging system that we calibrate according to the method proposed in [8]. ${}^b\mathbf{M}_o$ represents the constant homogeneous matrix between the robot base frame $\{R_b\}$ and the object frame $\{R_o\}$. It was roughly estimated by positioning the US probe plane such it crosses the middle of the object and such that the probe z-axis seems to be collinear with the object z-axis one. The parameters a_1 , a_2 and a_3 were also roughly estimated from the US image.

C. Control law

We use a very classical control law given by [12]:

$$\mathbf{v}_c = -\lambda \hat{\mathbf{L}}_s^+ (\mathbf{s} - \mathbf{s}^*) \quad (28)$$

where \mathbf{v}_c is the US probe instantaneous velocity sent to the low-level robot controller, λ is a positive gain, \mathbf{s}^* is the desired visual features vector, and $\hat{\mathbf{L}}_s^+$ is the pseudo inverse of the estimated interaction matrix $\hat{\mathbf{L}}_s$ given by:

$$\hat{\mathbf{L}}_s^+ = \hat{\mathbf{L}}_s^\top (\hat{\mathbf{L}}_s \hat{\mathbf{L}}_s^\top)^{-1} \quad (29)$$

IV. RESULTS

In both simulations and experiments the image moments were computed from the image coordinates of the points lying in the ultrasound cross-section image edge. For the simulations of section IV-B and *in-vitro* experiments, these image points were extracted by the use of a robust active snake algorithm. Image processing and control law computation were performed in real time at 25 frames/second thanks to the use of a PC computer equipped with a 3 GHz Dual core Xeon Intel processor running Linux.

A. Simulation results on a mathematical model

In a first part, we designed a simulator in the C++ language where the interaction between the probe and a perfect ellipsoid-shaped object is fully mathematically modelled in order to demonstrate the validity of the theoretical developments of this paper. For this simulation, images points of the section edge are computed directly from the mathematical interaction model. First, we test the case when the object parameters and its pose with respect to the US probe are assumed to be perfectly known. The object parameters are set to their exact values $(a_1, a_2, a_3) = (1, 2.5, 4)$ cm and

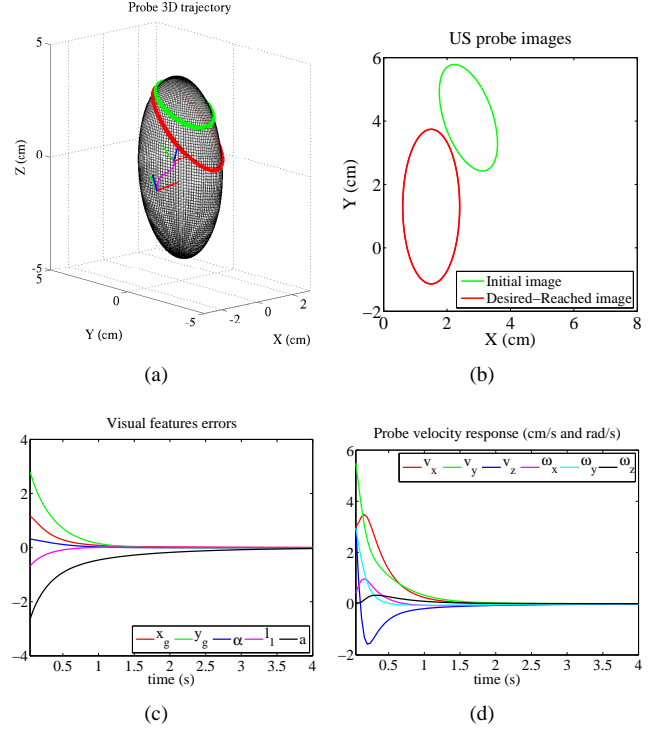


Fig. 3. Simulation results obtained for exact modeling parameters: (a) Probe 3D trajectory - (b) Initial (green) and desired-reached (red) image sections edges - (c) Visual error response (cm, cm, rad, cm and cm²) - (d) Velocity applied to the probe

the control gain λ is fixed to 1.5. The simulation results are shown in Fig. 3. The reached image (red) corresponds to the desired one (Fig. 3(b)) and the visual features errors converge exponentially to zero (Fig. 3(c)), thus demonstrating the validity of our theoretical developments. Starting from different initial poses, the pose reached by the probe does not always correspond to the pose where the desired features were learned since, as we mentioned before, there are several probe poses that give the same 2D ellipse in the image.

In a second test, we consider estimation errors on the object modeling parameters and pose. The parameters estimation errors are set to 50% for a_1 , a_2 and a_3 . The orientation and position estimation errors are set to 30 deg between $\{R_s\}$ and $\{R_o\}$ and 1 cm error on each of the three axes. The simulation results are shown in Fig. 4. Despite the large errors introduced, the visual features errors still converge to zero, thus demonstrating the robustness of the developed control system.

B. Simulation results with realistic US images

In a second part we use the ultrasound software simulator that was used in [13], which provides realistic US images of an egg-shaped object. This is useful for testing also the active snake algorithm to extract the section contour as it is necessary on real ultrasound image. This simulator allows to position and move a 2D virtual probe on a volume which is composed from 100 parallel real B-scan images. These US images have a resolution of 180×210 pixels with a pixel size of 0.2×0.2 mm and were previously captured from an

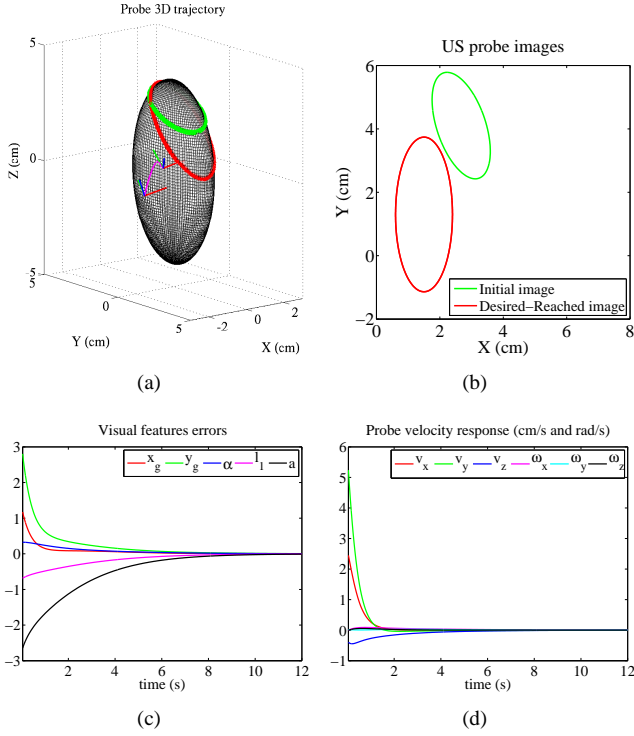


Fig. 4. Simulation results obtained for object parameters and pose estimation errors: (a) Probe 3D trajectory - (b) Initial (green) and desired-reached (red) image sections edges - (c) Visual error response (cm, cm, rad, cm, and cm^2) - (d) Velocity applied to the probe

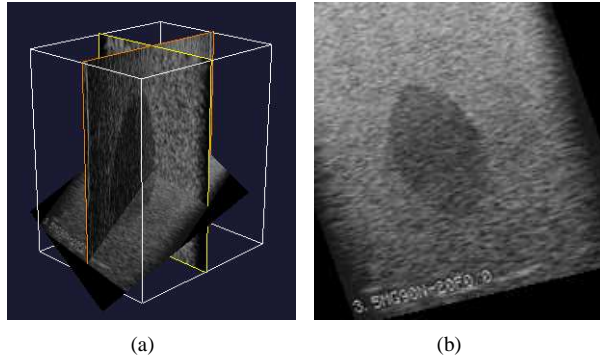


Fig. 5. The software simulator: (a) The virtual US probe interacting with the US volume - (b) The observed B-mode US image

ultrasound medical phantom at elevation intervals of 0.25 mm. The simulator is built from the Visualization Toolkit (VTK) software system [14] and our own ViSP library [15]. VTK is used to render the 3D view of the ultrasound volume and to generate the current 2D ultrasound image observed by the virtual probe by means of cubic interpolation, as displayed in Fig. 5(a) and Fig. 5(b) respectively. ViSP is used to compute the 2D moments from each virtual image and the visual servoing control law that is applied to the probe velocity. The 3D pose of the object frame with respect to the probe frame and its parameters were roughly estimated to ${}^b\mathbf{p}_o = (t_x, t_y, t_z, \theta_{u_x}, \theta_{u_y}, \theta_{u_z}) = (1.72, 2.69, 2.4, 0, 0, 0)$ (cm, deg) and $(a_1, a_2, a_3) = (0.56, 0.8, 1.65)$ cm respectively. Note that (t_x, t_y, t_z) represent the translational

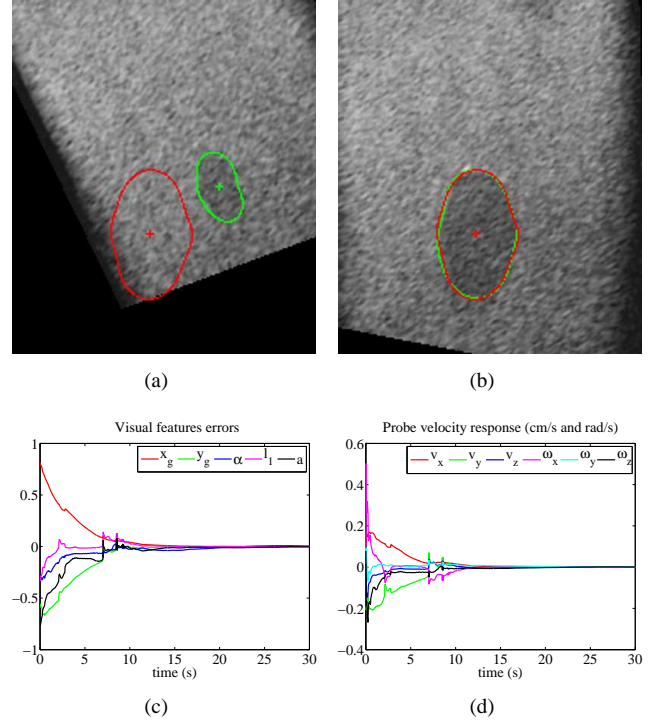


Fig. 6. Results from the ultrasound simulator: (a) Initial (green) and desired (red) image sections before applying visual servoing - (b) Current (green) and desired (red) after applying visual servoing (c) Visual error response (cm, cm, rad, cm and cm^2) - (d) Velocity applied to the probe

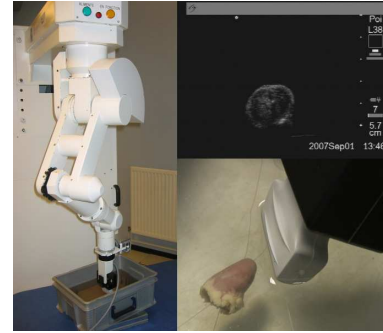


Fig. 7. Experimental setup: (left) ultrasound probe mounted on a 6 DOF medical robot - (right) Observed US image and the rabbit heart suspended in a water-filled box

components and $(\theta_{u_x}, \theta_{u_y}, \theta_{u_z})$ define the angles of the $\theta_{\mathbf{u}}$ representation. The control gain λ was set to 0.3. The results are depicted in Fig. 6. The visual features errors converge to zero roughly exponentially. This validates the method on an object having a different shape than an exact ellipsoid and therefore shows the robustness to object modeling error and measurement perturbations due to discontinuity in the edge detection.

C. In-vitro experimental results

Finally we test the method during *in-vitro* experiments where a 6-DOF medical robot arm similar to the Hippocrate system [16] is used to actuate a 5-10 MHz linear ultrasound

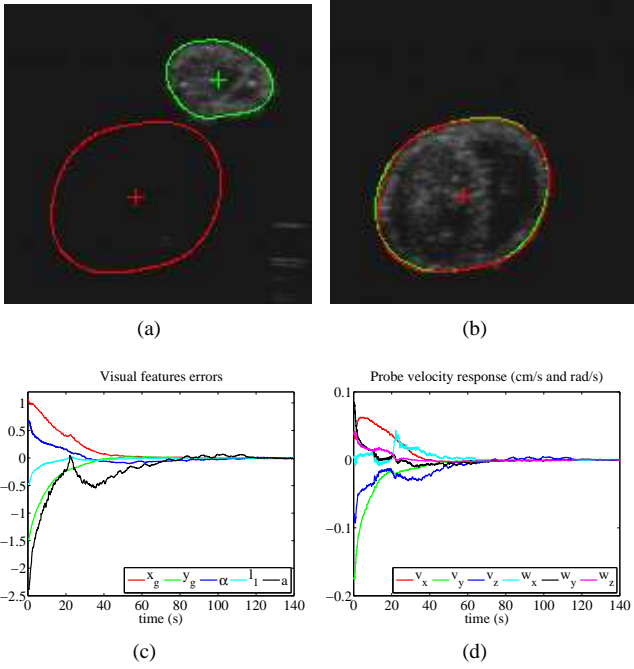


Fig. 8. Results from *in-vitro* experiment: (a) Initial (green) and desired (red) image sections edges before applying visual servoing - (b) Current (green) and desired (red) after applying visual servoing (c) Visual error response (cm, cm, rad, cm and cm^2) - (d) Velocity applied to the probe

probe transducer as shown in Fig. 7. The experiment² is performed using a rabbit heart suspended in a water-filled box by two nylon yarns. The gain λ is set to 0.07. The 3D pose of the object with respect to the robot base frame and the object axis length were roughly estimated respectively to ${}^b\mathbf{p}_o = (t_x, t_y, t_z, \theta u_x, \theta u_y, \theta u_z) = (0.10, 0.11, 0.84, 83, 23, 4)$ (cm, deg) and $(a_1, a_2, a_3) = (1.3, 1, 1.8)$ cm. The experimental results are depicted in Fig. 8. The visual features errors converge to zero and the reached ultrasound image corresponds to the desired one as can be seen in Fig. 8(b). During this experiment, we have noticed that the active snake allowing to extract the image section edge was very shaky due to the ultrasound noise. Nevertheless, in spite of this measurement noise and the difference between the shape of the rabbit heart and the theoretical ellipsoidal model considered in the modeling part, the proposed visual servoing succeeds with correct behaviour. Moreover the method has the advantage to be very robust to large initial error in the image as we can see in Fig. 8(b)-(a) where initial features are very far from the desired ones.

V. CONCLUSIONS AND FUTURE WORKS

This paper has presented a new ultrasound visual servoing based on image moments. First, the analytical form of the interaction matrix related to the ultrasound image moments was developed in the case of US probe interaction with an ellipsoidal object. Then, a set of visual features based on the combination of these moments was selected. Visual servoing was performed in both simulations and *in-vitro* experiments.

Successful results obtained with an ellipsoid object, an egg-shaped object and a real rabbit heart demonstrated the validity of our approach and its robustness with respect to modeling and measurements errors. Future work will concern the improvements of the method for more complex-shaped objects with no symmetry at all. We will also investigate on the optimal combination of moments which will result in a high interaction matrix decoupling for better performance. The great challenge is to develop a generic interaction matrix for any kind of object that requires the least modeling parameters as possible.

ACKNOWLEDGMENTS

The authors acknowledge Christophe Collewet for providing the image processing algorithm and Seth Hutchinson for his comments on this research and earlier draft of this paper.

REFERENCES

- [1] P. Abolmaesumi, S. E. Salcudean, W-H Zhu, M. R. Sirouspour, and S P. DiMaio. Image-guided control of a robot for medical ultrasound. *IEEE Trans. on Robotics*, vol. 18, no. 1, february 2002.
- [2] T. R. Nelson and D. H. Pretorius. Three-dimensional ultrasound imaging. *Ultrasound in Medicine and Biol*, 24(9):1243-1270, 1998.
- [3] J. Hong, T. Dohi, M. Hashizume, K. Konishi, N. Hata, An ultrasound-driven needle insertion robot for percutaneous cholecystostomy, *Physics in Medicine and Biology*, 49(3):441-445, 2004.
- [4] J. Stoll, P. Novotny, R. Howe, P. Dupont, Real-time 3D ultrasound based servoing for a surgical instrument, in *IEEE Int. Conf. on Robotics and Automation*, Orlando, Florida, May 2006.
- [5] P.M. Novotny, J.A. Stoll, P.E. Dupont, R.D. Howe. Real-time visual servoing of a robot using three-dimensional ultrasound, in *IEEE Int. Conf. on Robotics and Automation*, p. 2655-2660, Roma, Italia, May 2007.
- [6] M. A. Vitrani, G. Morel, N. Bonnet, and M. Karouia. A robust ultrasound-based visual servoing approach for automatic guidance of a surgical instrument with *in vivo* experiments. *Biomedical Robotics and Biomechanics BioRob, The first IEEE RAS/EMBS*, pp. 35-40, Spain, february, 2006.
- [7] M. A. Vitrani, H. Mitterhofer, N Bonnet, G. Morel. Robust ultrasound-based visual servoing for beating heart intracardiac surgery, in *IEEE Int. Conf. on Robotics and Automation*, p. 3021-3027, Roma, Italia, May 2007.
- [8] A. Krupa. Automatic calibration of a robotized 3D ultrasound imaging system by visual servoing. In *IEEE Int. Conf. on Robotics and Automation, ICRA'2006*, p. 4136-4141, Orlando, Florida, May 2006.
- [9] W. Bacht and A. Krupa. Towards ultrasound image-based visual servoing. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pp. 4112-4117, Orlando, Florida, May 2006.
- [10] A. G. Mamistvalov. N-dimensional invariants and conceptual mathematical theory of recognition n-dimensional solids. In *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20 no. 8, august 1998.
- [11] F. Chaumette. Image moments: A general and useful set of features for visual servoing. *IEEE Trans. on Robotics*, 20(4):713-723, 2004.
- [12] B. Espiau, F. Chaumette and P. Rives. A new approach to visual servoing in robotics. in *IEEE Trans. on Robotics*, 8(3):313-326, June 1992.
- [13] A. Krupa, G. Fichtinger and G. D. Hager. Full motion tracking in ultrasound using image speckle information and visual servoing. In *IEEE int. Conf. on Robotics and Automation ICRA'07*, p. 2458-2464, Roma, Italia, May 2007.
- [14] W. Schroeder, K. Martin, B. Lorensen. *The Visualization Toolkit An Object Oriented Approach To 3D Graphics*, 3rd Edition, ISBN, 1-930934-12-2 Kitware, Inc. Publishers.
- [15] E. Marchand, F. Spindler, F. Chaumette. ViSP for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics and Automation Magazine*, 12(4):40-52, 2005.
- [16] F. Pierrot, E. Dombre, E. Degoulange, L. Urbain, P. Caron, S. Boudet, J. Gariepy, and J. Megnien, Hippocrate: A safe robot arm for medical applications with force feedback, *Medical Image Analysis (MedIA)*, vol. 3, no. 3, pp. 285-300, 1999.

²A short video of the experiment accompanies the paper.